


Perceptual Evaluation of Signal-to-Noise-Ratio-Aware Dynamic Range Compression in Hearing Aids

Trends in Hearing
Volume 24: 1–14
© The Author(s) 2020
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/2331216520930531
journals.sagepub.com/home/tia


Borys Kowalewski , Torsten Dau, and Tobias May 

Abstract

Dynamic range compression is a compensation strategy commonly used in modern hearing aids. Fast-acting systems respond relatively quickly to the fluctuations in the input level. This allows for more effective compression of the dynamic range of speech and hence enhanced the audibility of its low-intensity components. However, such processing also amplifies the background noise, distorts the modulation spectra of both the speech and the background, and can reduce the output signal-to-noise ratio (SNR). Recently, May et al. proposed a novel SNR-aware compression strategy, in which the compression speed is adapted depending on whether speech is present or absent. Fast-acting compression is applied to speech-dominated time–frequency (T-F) units, while noise-dominated T-F units are processed using slow-acting compression. It has been shown that this strategy provides a similar effective compression of the speech dynamic range as conventional fast-acting compression, while introducing fewer distortions of the modulation spectrum of the background and providing an improved output SNR. In this study, this SNR-aware compression strategy was compared with conventional fast- and slow-acting compression in terms of speech intelligibility and subjective preference in a group of 17 hearing-impaired listeners with varying degree of hearing loss. The results show a speech intelligibility benefit of the SNR-aware compression strategy over the conventional slow-acting system. Furthermore, the SNR-aware approach demonstrates an increased subjective preference compared with both conventional fast- and slow-acting systems.

Keywords

wide dynamic range compression, signal-to-noise ratio, hearing-aid signal processing, speech intelligibility, subjective preference

Received 14 August 2019; Revised 7 May 2020; accepted 8 May 2020

Sensorineural hearing loss is associated with a decreased sensitivity to low-intensity sounds as well as a range of suprathreshold auditory deficits. These deficits include, among others, the phenomenon of loudness recruitment and the limitation of the dynamic range (e.g., Bacon & Oxenham, 2004; Smeds & Leijon, 2011). To account for this, modern hearing aids typically implement some form of level-dependent amplification such as wide dynamic range compression (WDRC, see Souza, 2002, for a review). Such systems provide relatively high gain for low-intensity input sounds to ensure sufficient audibility, which appears to be necessary for good speech recognition (Pavlovic & Studebaker, 1984; Souza & Turner, 1999; Woods et al., 2013). As the input level increases, the gain is reduced to avoid loudness discomfort. To follow the temporal dynamics of speech, a compression

system should respond rapidly to changes in the input level across time (Edwards, 2004; Moore, 2008; Souza, 2002). This requires the use of short time constants in the level estimation stage of the signal-processing chain (for implementation details, see Giannoulis et al., 2012; Kates, 1993). However, the application of short time constants can also lead to rapid fluctuations in the gain function over time, introducing potentially detrimental distortions of the temporal envelope of speech

Hearing Systems Section, Department of Health Technology, Technical University of Denmark

Corresponding Author:

Tobias May, Hearing Systems Section, Department of Health Technology, Technical University of Denmark, Kgs. Lyngby, Denmark.
Email: tobmay@dtu.dk



(e.g., Gatehouse et al., 2006; Jenstad & Souza, 2005, 2007; Plomp, 1988; Souza et al., 2012a; Walaszek, 2008). A number of studies have shown that fast-acting WDRC provides an improvement in audibility of speech sufficient to offset the potentially detrimental distortion of the temporal envelope of the signal, leading to a net intelligibility benefit. This was demonstrated for speech in quiet by Villchur (1973), Souza and Turner (1998, 1999), Souza and Bishop (1999), and Davies-Venn et al. (2009). An acoustic analysis conducted by Alexander and Rallapalli (2017) showed that fast-acting compression leads to a higher effective compression ratio (ECR, based on short-term level histograms¹) compared with slow-acting compression. This has a positive effect on speech audibility but, on the other hand, negatively affects the speech modulation transfer function (MTF). The speech-recognition results reported in the same study suggest that, in many cases, the audibility benefit counteracts the negative effects of envelope-domain distortion.

When the target speech is degraded by background noise, the benefit of WDRC appears to depend on a variety of factors such as the spectrotemporal characteristics of the noise, the overall input level, and the signal-to-noise ratio (SNR), as demonstrated, for example, by Yund and Buckles (1995). Souza et al. (2006) demonstrated that the presence of background noise decreases the overall amount of envelope fluctuations, leading to less dynamic changes in the gain function and, as a result, a decreased ECR of speech. Rhebergen et al. (2009) reported beneficial effects of compression on the speech reception threshold (SRT) when the processing was applied to the speech alone prior to mixing it with the background noise. However, such conditions are rather artificial. Rhebergen et al. considered also a more realistic scenario, in which the processing was instead applied to the mixture of speech and either a stationary or a nonstationary, interrupted noise. In that case, compression had a pronounced beneficial effect on the SRT in the interrupted noise. Similar findings were reported in a later study by Rhebergen et al. (2017). At negative SNRs (as was the case in both studies of Rhebergen et al.), the interferer is the more dominant stimulus and its temporal fluctuations drive the compression system. The gain is increased during the dips in the noise, amplifying the low-level glimpses of speech present in those dips. The results of Desloge et al. (2017) and Kowalewski et al. (2018) further support the notion that fast-acting compression systems provide improved short-term audibility and increased opportunities for glimpsing, as long as the noise exhibits prominent fluctuations and the long-term input SNR is negative.

In contrast, in scenarios characterized by high long-term input SNRs, the compression is driven mostly by the changes of the speech level. The fast changes in gain

cyclically amplify the background, introducing modulation components to the noise (Stone & Moore 2003, 2004, 2007, 2008) and reducing the long-term output SNR (Hagerman & Olofsson, 2004; May et al., 2018; Naylor & Johannesson, 2009; Rhebergen et al., 2009, 2017; Souza et al., 2006). Both effects are potentially detrimental to speech intelligibility and the perceived sound quality. Taken together, the previous findings indicate that fast-acting compression has rather positive effects on speech intelligibility due to increased audibility and a reduced dynamic range in the following scenarios: (a) speech in quiet, (b) speech in the presence of a strongly fluctuating noise at a negative SNR, and (c) speech compressed prior to mixing it with noise (unrealistic). These benefits are largely reduced, or turn into a detriment, as soon as the input SNR becomes positive (which is a common scenario, see Smeds et al., 2015; Weisser & Buchholz, 2019) and/or when the interferer is stationary. It is nevertheless possible that the advantages of fast-acting compression would be restored if a selective processing of the speech and the noise components could be achieved.

Several studies have focused on the effects of compression release time on listener's subjective preference and/or perceived quality. Their conclusions are largely in line with the aforementioned studies on speech intelligibility. Neuman et al. (1995) investigated hearing-impaired (HI) listeners' overall preference for the compression release time (60, 200, and 1000 ms) when processing speech in the presence of background noise of varying characteristics and levels. Overall, longer release times were preferred for the types of noise naturally characterized by higher sound pressure levels (SPLs). In a follow-up study using the same set of conditions (Neuman et al., 1998), the listeners rated several attributes of sound quality. The results indicated that, with longer compression release times, the ratings of the overall impression, pleasantness, and clarity increased, while the rating of noisiness decreased. This was likely due to the above-mentioned cyclical amplification of the background noise that occurs at positive input SNRs. The effect becomes more prominent with shorter release times (as more gain is provided to the noise during the speech gaps) and is more noticeable as the level of the background increases. A similar preference for longer release times was demonstrated by Hansen (2002) in a group of HI listeners and a range of acoustic scenarios. Neuman et al. (1995) suggested the use of an adaptive release time in hearing aids in order to improve the perceived sound quality. A shorter release time could be used in quieter scenarios, while a longer release time could be applied with increasing levels of background noise. Several adaptive compression strategies have been proposed in the past including the K-AMP (Killion et al., 1992), the dual front-end adaptive gain

control (Moore & Glasberg, 1988), the guided level estimator (Neumann, 2008), and the short-term dynamic-range-driven system proposed by Lai et al. (2013). However, all of these systems rely on short-term level dynamics of the speech and noise mixture and do not explicitly utilize information related to the presence of the target signal with respect to the background noise.

The SNR-aware dynamic range compression strategy presented by May et al. (2018) attempts to combine the advantages of both fast- and slow-acting compression. The main idea is to adjust the release time of the compressor in each individual time–frequency (T-F) unit depending on whether the target is present or absent. Specifically, a short release time is applied to speech-dominated T-F units where the short-term SNR is high, while a longer release time is used to process T-F units that are dominated by noise. The SNR-aware compression strategy bears some similarities with the aforementioned artificially created scenario tested by Rhebergen et al. (2009), where the speech alone was compressed prior to mixing it with noise. The difference is that the SNR-aware approach operates on the noisy speech mixture and does not require the availability of separate speech and noise signals, making it potentially applicable in hearing devices. Similar principles had previously been applied in the compression system driven by the direct-to-reverberant energy ratio, which was shown to preserve the listeners' spatial perception (Hassager et al., 2017). May et al. (2018) provided an instrumental evaluation of the SNR-aware compression strategy and compared it with conventional fast- and slow-acting compression. The SNR-aware compression strategy provided ECRs similar to those obtained with conventional fast-acting compression, while the natural fluctuations in the background noise were preserved in a similar way as when slow-acting compression was applied.

In this study, the SNR-aware dynamic range compression strategy was evaluated in terms of speech intelligibility and subjective preference in a group of HI listeners. It was hypothesized that the SNR-aware compression strategy would provide superior audibility compared with slow-acting compression, while it would result in a higher output SNR and introduce fewer distortions of the background compared with fast-acting compression, leading to superior speech intelligibility performance and higher preference scores. To exclude the potential effects of SNR estimation errors on perception, the ideal SNR-aware strategy based on the *a priori* SNR was tested.

Methods

Participants

The study included 17 HI listeners aged 25 to 80 years (average 68.7 years). All participants underwent

screening conducted by a trained audiologist, which included tympanometry, pure-tone audiometry (air and bone conduction), and word recognition scores in quiet (discrimination scores) using the Dantale corpus (Elberling et al., 1989). Based on this evaluation, all listeners' hearing loss was classified as sensorineural. The listeners' audiograms were compared with the standard audiograms proposed by Bisgaard et al. (2010) and were further classified into three groups based on the smallest absolute distance criterion (in dB): seven listeners in the N_2 group, seven listeners in the N_3 group, and three listeners in the N_4 group. The tested ear was chosen based on the best match to the desired hearing profile and/or the best discrimination score. To ensure that an SRT could be reliably measured in noise, a discrimination score exceeding 80% was required as an inclusion criterion. The listeners' audiograms are shown in Figure 1. All listeners were native speakers of the Danish language. After a short introduction to the test procedure, they provided an informed consent. All listening tests were conducted at the Technical University of Denmark. The experiments were approved by the Science-Ethics Committee for the Capital Region of Denmark (Reference H-16036391).

Signal Processing and Fitting

All dynamic range compression systems were based on the short-time discrete Fourier transform using frames of 10 ms duration with 75% overlap and operated in seven independent octave-wide frequency channels with center frequencies ranging from 125 to 8000 Hz. The level estimation in each frequency channel was performed using a first-order infinite impulse response filter with different time constants associated with the attack and the release (Kates, 1993). As shown in Table 1, the following three compression systems were tested: conventional fast- and slow-acting compression as well as SNR-aware compression. The attack time in the level estimator was always set to 5 ms. The fast-acting system utilized a level estimator with a short release time of 40 ms, while it was set to 2000 ms for the slow-acting system. The level estimator in the SNR-aware system switched between the short and the long release time in individual T-F units by applying a threshold criterion of 0 dB to the *a priori* SNR. If the *a priori* SNR was higher than 0 dB, the corresponding T-F unit was processed with the short release time, resulting in a fast-acting system. Otherwise, if the *a priori* SNR was lower than 0 dB, the long release time was used, resulting in a slow-acting system. The *a priori* SNR was calculated by comparing the energy of the separate speech and noise signals in individual T-F units. More details concerning the implementation of the algorithms can be found in May et al. (2018).

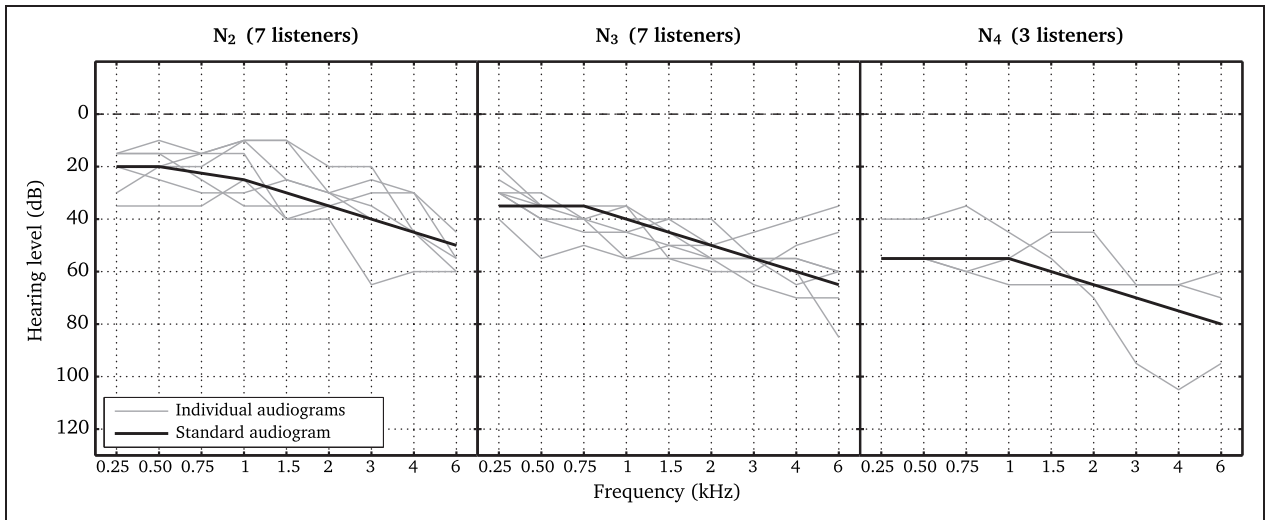


Figure 1. Pure-Tone Audiograms of Listeners in the N_2 , N_3 , and N_4 Groups. Individual audiograms of all listeners in a given group are shown with a gray line, while the corresponding standard audiogram is indicated by the thick black line. The audiograms are shown for frequencies up to 6 kHz, which is the uppermost frequency in the profiles provided by Bisgaard et al. (2010).

Table 1. Configuration of the Three Tested Compression Schemes.

Compressor	Attack (ms)	Release (ms)	Speech detection	Estimator
Fast	5	40	off	—
Slow	5	2000	off	—
SNR-aware	5/5	40/2000	on	<i>a priori</i> SNR

Note. SNR = signal-to-noise ratio.

The compression thresholds (CTs) in each frequency channel were calibrated using a stationary noise with an SPL of 50 dB and a spectrum that was spectrally matched to the long-term average spectrum of the Danish hearing-in-noise test (HINT) corpus. Linear (level-independent) gain was applied below the CT. The linear gain and compression ratios (CRs) were calculated from the insertion gain for 50 and 80 dB SPL prescribed by the National Acoustic Laboratories Non-Linear 2 (NAL-NL2; Keidser et al., 2011) rationale. In the fitting software, the settings *unilateral* and *slow* were selected. The former setting was chosen to take the monaural presentation of the stimuli into account. The latter setting was chosen because the NAL-NL2 rationale provides higher nominal CRs for slow-acting compression (Keidser et al., 2011), which should further increase the acoustic differences between the processing conditions. To reduce the inter-listener variability of the compression parameters, the CRs were fitted on a group level. The CRs for each group of listeners were based on the fitting to the respective standard audiograms (i.e., N_2 , N_3 , or N_4). Table 2 shows the CTs and the CRs

for individual frequency channels. The linear gain, on the other hand, was fitted individually for each listener for the sake of audibility of the stimulus portions that fall below the CT.

Stimuli and Procedure

Noisy speech sampled at a rate of 20 kHz was created by mixing clean speech from the Danish HINT corpus (Nielsen & Dau, 2011) with the following two noise types: the stationary International Collegium of Rehabilitative Audiology (ICRA)-1 noise (Dreschler et al., 2001) and the factory noise from the NOISEX database (Varga & Steeneken, 1993). The factory noise was a recording from an industrial production plant, consisting of various acoustic events, including machine and conveyor belt sounds, with a moderate degree of reverberation. It therefore contained natural spectrotemporal fluctuations, in contrast to the stationary background (which only contained intrinsic temporal fluctuations). The two noise types were chosen in order to investigate potential perceptual effects of spectrotemporal interactions between speech and the background. Both were spectrally matched to the long-term average spectrum of the HINT corpus measured in one-third octave bands. For each noisy speech mixture, a random noise segment was selected. A noise-only segment of 1 s duration was included before and after each sentence.

The administration of the tests and the preprocessing of stimuli were performed using a personal computer running MATLAB. The stimuli were delivered through an RME Fireface UCX audio interface with a 16-bit

Table 2. Compression Thresholds (CTs) in dB and Compression Ratios (CRs) for Individual Channel Center Frequencies.

		Channel center frequency (Hz)						
		125	250	500	1000	2000	4000	8000
CT (dB)		43	43	41	41	37	31	28
CR	N_2	1.1:1	1.1:1	1.2:1	1.4:1	1.8:1	2.2:1	2.0:1
	N_3	1.7:1	1.7:1	1.8:1	2.1:1	2.6:1	2.8:1	2.4:1
	N_4	2.2:1	2.2:1	2.2:1	3.0:1	3.5:1	3.3:1	2.5:1

resolution, connected to an SPL Phonitor Mini head-phone preamplifier. The listeners were placed in a double-walled soundproof booth and listened monaurally through Sennheiser HDA200 headphones. The SPL of the stimuli was calibrated using a Brüel & Kjær Type 4153 ear simulator (IEC, 2009). The frequency response of the headphones was measured in the same simulator and equalized using a digital filter to have a flat response in the coupler microphone.² The SPL of the noise was fixed at 50 dB. The noise level was chosen such that the entire speech and noise mixture were at a relatively low SPL, emphasizing the influence of audibility on speech perception. During the speech intelligibility test, a trained audiologist (a native speaker of Danish) was present in the booth with the listeners and performed the scoring on a computer screen. The paired-comparison preference judgment was executed with the participants seated in front of the screen themselves and providing responses using a graphical user interface. Before each part of the experiment, the listeners were given spoken instructions regarding the procedure.

SRT Determination. The experimental session began with measuring the SRT in each noise type using conventional fast-acting compression. Scoring was performed on a sentence basis, that is, a correct recall of all five words was required to mark the presented sentence as correct. Each list consisted of 20 sentences. Following the listener's response to each sentence, the SNR for the next sentence was determined and stored (also following the last sentence on the list, yielding 21 stored SNRs, Nielsen & Dau, 2011). The start SNR was +5 dB. If the first sentence was not correctly identified, it was repeated with an increasing SNR until recalled correctly. The initial step size was 4 dB and was reduced to 2 dB after the first five sentences (Nielsen & Dau, 2011). The SRT was determined as the average of the SNRs from sentence 6 to 21. For each noise type, a training trial was conducted using the HINT training lists. Subsequently, two estimates of the SRT were made (test trials) using an HINT test list selected at random (without replacement). The final SRT value for each noise type was determined as a mean of the values obtained using the two test lists. The starting noise type was selected at random and the noise types were subsequently alternated.

Fixed-SNR Sentence-Recognition Scores. A sentence-recognition score was determined for each of the six conditions (2 Noise Types \times 3 Processing Strategies). The SNR was fixed for each noise type and equal to the corresponding SRT, determined in the first part of the experiment. The order of the conditions was randomized for each listener. However, each test list was immediately preceded by a training list in the corresponding condition, in order to familiarize the listeners with the given combination of noise and processing type over a broad range of SNRs. The six HINT test lists remaining after SRT determination were selected at random (without replacement). The training lists were used with replacement, such that some of the training lists were experienced by the listeners multiple times in different conditions throughout the entire experiment.

Paired-Comparison Preference Test. For each of the two noise types, comparisons between all three processing types were made (six comparisons in total). Each listener completed 3 trials, for a total of 18 comparisons (except for 1 participant who completed only 2 trials or 12 comparisons).

Before each trial, three sentences from the HINT corpus were selected at random and concatenated to create a running speech sample. The sample was mixed with the background noise at the same SNR as used in the preceding sentence-recognition test. In each presentation, the speech-in-noise sample was processed with each of the processing strategies and presented to the listeners as *interval A* and *interval B* in a two-alternative forced-choice manner. The question displayed on the screen was "Which interval do you prefer?". The order of the processing conditions and noise types was randomized within a trial. The assignment of processing conditions to *A* or *B* was also randomized within each presentation. The listeners were blinded to the processing conditions being compared. They had to listen to the entire length of *A* and *B* prior to indicating their preference. They could also repeat each of the intervals separately as many times as needed. The listeners were instructed to base their decisions on subjective judgments of overall sound quality and to pay attention to such attributes as quality of the speech, subjective intelligibility, characteristics of the noise or listening

comfort, but not to focus on one single attribute in particular. If there was no perceived difference between the intervals, the listeners were instructed to pick an interval at random.

Results

Speech reception thresholds

The individual SRTs are shown in Figure 2 for each noise type as a function of the hearing profile (N_2 , N_3 , and N_4). A two-way fixed-factor analysis of variance (ANOVA) was conducted on the SRT data, with factors *noise type* and *hearing profile*. It has to be noted that the results of the analysis should be interpreted carefully, as

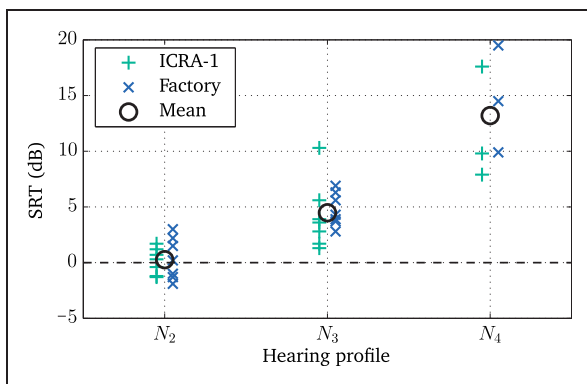


Figure 2. Individual SRTs of all Listeners for ICRA-I and Factory Noise as a Function of the Hearing Profile (N_2 , N_3 , and N_4). The mean SRT averaged across noise types and all listeners within a hearing profile are shown by the black circles. ICRA-I = International Collegium of Rehabilitative Audiology Souza; SRT = speech reception threshold.

the N_4 group included a smaller number of participants than the other two groups (three vs. seven). On the group level, the results indicated a significant main effect of hearing profile ($F = 49.71$, $df = 2$, $p < .001$) and no effect of noise type, nor any significant interaction between noise type and hearing profile. The SRT averaged across noise types and all listeners within a hearing profile was 0.26, 4.48, and 13.20 dB for the N_2 , N_3 , and N_4 groups, respectively. These mean SRT values are indicated by the black circles in Figure 2.

Sentence Scores

A rationalized arcsine units (RAUs) transform (Studebaker, 1985) was applied to the sentence-recognition scores expressed in percent correct. The RAU-transformed scores were averaged across listeners and are shown in Figure 3 as a function of the processing type (fast, slow, and SNR-aware compression) for the ICRA-1 noise (left panel) and the factory noise (right panel). Subsequently, a three-way, mixed-effects ANOVA was conducted on the transformed data. The fixed factors were *noise type* (with two levels: ICRA-1 and factory), *processing type* (with three levels: fast, slow, and SNR-aware) and *listener*. The listener was included as a random factor to account for the variability in the degree of hearing loss, differences in audibility, sensitivity to distortion, the operating SNR, and so on (Naylor, 2016). In addition, all possible first-order interactions were included.

The ANOVA revealed a large and significant main effect of *processing type* ($F = 4.07$, $df = 2$, $p = .0266$, partial $\eta^2 = 0.21$, Cohen, 1973). Moreover, a significant interaction between the *noise type* and *listener* was found ($F = 2.84$, $df = 16$, $p = .0059$, partial $\eta^2 = 0.59$). The

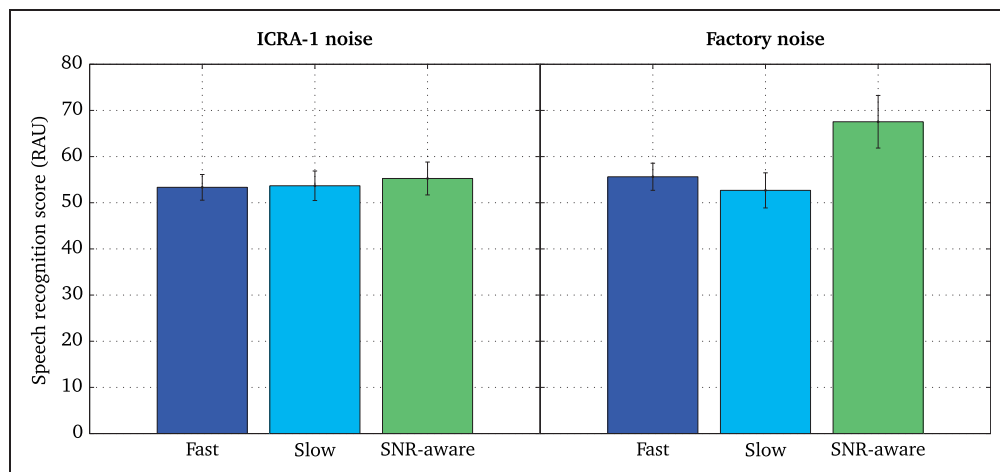


Figure 3. RAU-Transformed Sentence Recognition Scores Averaged Across Listeners as a Function of the Processing Type (Fast, Slow, and SNR-Aware Dynamic Range Compression) for ICRA-1 Noise (Left Panel) and Factory Noise (Right Panel). The error bars indicate the standard errors of the mean. RAU = rationalized arcsine units; SNR = signal-to-noise ratio; ICRA-1 = International Collegium of Rehabilitative Audiology.

interaction between the factors *noise type* and *processing type* did not reach statistical significance ($F=2.61$, $df=2$, $p=.089$). Therefore, in the post hoc analysis, the results were pooled across both noise types. For each processing type, the RAU-transformed scores were averaged across listeners, as shown in Figure 4.

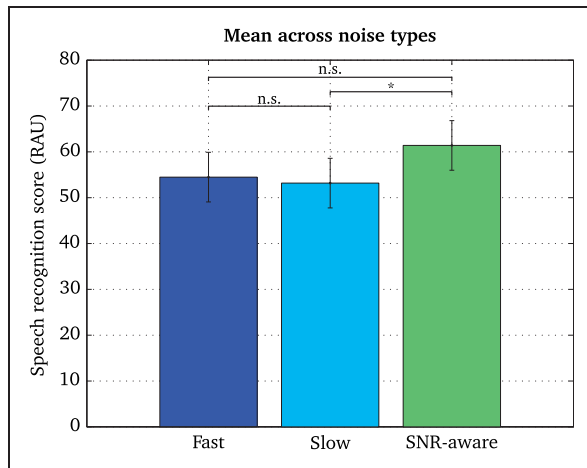


Figure 4. RAU-Transformed Sentence-Recognition Scores Averaged Across Listeners and Noise Types as a Function of the Processing Type (Fast-, Slow-, and SNR-Aware Dynamic Range Compression). The error bars represent the 95% confidence intervals (see the main text for details). Level of statistical significance of the difference of means is indicated as follows: * .05 or *ns* = nonsignificant. RAU = rationalized arcsine units; SNR = signal-to-noise ratio; ICRA-I = International Collegium of Rehabilitative Audiology.

For the sake of comparison of the means, 95% confidence intervals were constructed based on the mean squared error from the ANOVA and their lengths were adjusted using Bonferroni corrections to account for multiple comparisons. The post hoc analysis revealed no statistically significant differences between the fast and the SNR-aware system, nor between the fast and the slow system. The only statistically significant difference was found between the slow and the SNR-aware system (61.4 vs. 53.2 RAU, $p < .05$).

Subjective Preference

For each noise type, data from 150 paired-comparison trials were collected (16 Listeners \times 9 Trials + 1 Listener \times 6 Trials). For each listener, the trials were evaluated for consistency in terms of transitivity, and the trials containing circular triads were rejected³(see Kendall, 1962; Kendall & Smith, 1940, for a detailed discussion). Overall, 111 of the 150 trials for the ICRA-1 noise and 120 of the 150 trials for the factory noise were considered for further analysis. For each noise type, the responses from the remaining trials were pooled together to create response matrices. These matrices are summarized in terms of the number of wins for each strategy in the top panels of Figure 5. Subsequently, the values in the response matrices were converted to relative frequency and evaluated for weak stochastic transitivity⁴ (Ellermeier et al., 2004). The weak stochastic transitivity was maintained for both noise types, which allowed to fit a more restrictive

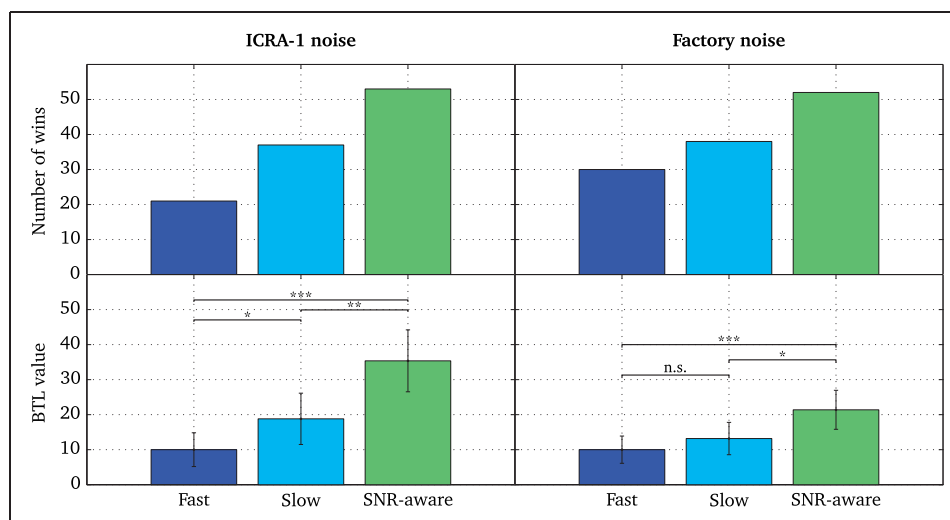


Figure 5. Results of the Subjective Preference Test as a Function of the Processing Type (Fast-, Slow-, and SNR-Aware Dynamic Range Compression) for ICRA-1 Noise (Left Panels) and Factory Noise (Right Panels). The panels in the top row show the number of wins based on the consistent trials from all listeners. The panels in the bottom row show the corresponding ratio-scale values derived from the BTL model, including the 95% confidence intervals (see the main text for details). Level of statistical significance is indicated as follows: * .05, ** .01, *** .001 or *ns* = nonsignificant; SNR = signal-to-noise ratio; BTL = Bradley-Terry-Luce; ICRA-I = International Collegium of Rehabilitative Audiology.

Bradley–Terry–Luce (BTL) model (Bradley & Terry 1952; Ellermeier et al., 2004; Luce, 1959). The BTL model was evaluated separately for each noise type using the MATLAB function provided by Wickelmaier and Schmid (2004). In either case, model validity could not be rejected as indicated by the likelihood-ratio test ($\chi^2 = 0.53$, $p = .46$ for ICRA-1, $\chi^2 = 0.02$, $p = .89$ for factory noise). The model output is represented by ratio-scale values that reflect how likely a given item is to be preferred in comparison with another, randomly selected one. These values are shown in the bottom panels of Figure 5, together with the corresponding 95% confidence intervals. In each case, the fast condition was chosen as a reference and arbitrarily assigned a value of 10. In the presence of ICRA-1 noise, significant differences were found between the BTL scale values for all pairwise comparisons of processing types (35.37, 18.81, and 10.00 for SNR-aware, slow, and fast, respectively). In the factory noise condition, the BTL value for the SNR-aware compression strategy was significantly higher than those for slow and fast compression (21.38 vs. 13.18 and 10.00), but there was no significant difference between slow and fast compression.

Discussion

The purpose of this study was to conduct a perceptual evaluation of the novel SNR-aware compression strategy proposed by May et al. (2018) in HI listeners. Three audiometrically profiled groups were tested: N_2 , N_3 , and N_4 . Two noise types were considered: ICRA-1 stationary speech-shaped noise and factory noise from the NOISEX database. The SNR-aware strategy was compared with conventional fast- and slow-acting compression systems. For each noise type, the listeners' individual SRTs were determined using fast-acting compression. The corresponding SNR values were subsequently used for obtaining sentence-recognition scores at a fixed SNR, as well as preference ratings using a paired-comparison paradigm.

Compression Strategy

The ANOVA of sentence-recognition scores indicated a statistically significant main effect of processing type and no main effect of noise type. Moreover, the interaction between the noise type and the processing type did not reach statistical significance. However, the following trend was observed in the RAU-transformed sentence-recognition scores shown in Figure 3. In the ICRA-1 noise, it appears that there are almost no differences between the (averaged) scores. While a small advantage of fast- versus slow-acting compression was found in the factory noise condition, a larger advantage over either of the two conventional schemes was obtained with the

SNR-aware processing scheme. Because the interaction was not statistically significant, the subsequent post hoc tests had to be conducted on scores pooled across noise types. Nevertheless, it appears that the pattern observed in the analysis might be *blurred* by the outcomes obtained with the ICRA-1 noise. The post hoc tests revealed an advantage of the SNR-aware strategy over conventional slow-acting compression and no difference between the SNR-aware and the conventional fast-acting processing.

Compared with slow-acting compression, fast-acting compression of speech provides ECRs that are closer to the nominal CR prescribed by the gain rationale, resulting in improved audibility. The results of this study suggest that these acoustic effects are necessary (but not sufficient) for an improved speech recognition in noise. If conventional processing is applied, those positive effects are likely offset by a distortion of the noise modulation spectrum and a reduction of the long-term broadband SNR. To take full advantage of fast-acting compression, a differentiation between the target and the background is required, followed by applying some distinct processing to the two signal components (foreground vs. background). This is achieved by the SNR-aware compression strategy and seems to provide a more favorable balance between audibility and ECR improvement versus MTF- and SNR-distortion. Moreover, as mentioned earlier, the advantage of the SNR-aware strategy seems to be more pronounced in the factory noise condition. This could stem from the stronger interaction between the speech and the background noise due to natural envelope fluctuations occurring in the two signals. The SNR-aware compression strategy reduces this interaction which could be advantageous for speech recognition. However, this interpretation has to be treated with caution due to the weak statistical evidence supporting it.

The subjective preference scores indicated an advantage of the novel SNR-aware compression strategy over both conventional fast- and slow-acting processing for both noise types. In addition, an advantage of slow- over fast-acting compression was observed in the stationary ICRA-1 noise but not in the nonstationary factory noise. This suggests that the cyclical amplification has a more prominent negative effect on the perceived quality in stationary backgrounds. This is consistent with the conclusion drawn by Neuman et al. (1995), that the cyclical *pumping* becomes more noticeable as more noise is present in the speech gaps. Informally, some of the participants in this study reported that most of the perceived differences between the compared strategies were in the characteristics of the background noise. The additional advantage of SNR-aware over slow-acting compression likely stems from the increased ECR and improved audibility, which are potentially linked to improved speech

intelligibility. It is likely that the listeners' ability to comprehend the processed speech material was an important factor that contributed to the overall preference judgment. This is consistent with the studies by Preminger and Van Tasell (1995) and Hansen (2002). Preminger and Van Tasell investigated the effects of different frequency shaping on normal-hearing listeners' ratings in terms of several attributes of subjective sound quality such as intelligibility, pleasantness, listening effort, loudness, and overall impression. They found that ratings across the other dimensions were correlated with the ratings of intelligibility. Hansen tested HI listeners' preference in terms of several attributes of sound, including subjective intelligibility using WDRC-processed stimuli with various combinations of time constants and CTs. The conditions yielding the highest overall preference also corresponded to the highest preference in terms of subjective intelligibility.

Listener-Specific Factors

As expected, the SRT depended on the degree of hearing loss and was highest (the worst) in the N_4 group, which is shown in Figure 2 and indicated by the ANOVA. The N_4 listeners were hence tested at the highest SNRs in the subsequent parts of the experiment. Therefore, they experienced greater acoustic differences between the processing strategies (May et al., 2018). This phenomenon was described by Naylor (2016) as *selection-treatment interaction*, that is, a situation in which the selection of the participants (their hearing profiles and therefore the SRTs) influences the magnitude of the differences across treatments (processing strategies) and was the main reason to include *listener* as a random factor in the statistical analysis of sentence recognition scores. As the listeners were tested in the vicinity of the steepest point on the psychometric function, large acoustic differences were, in turn, expected to create large perceptual differences. The beneficial effects of SNR-aware compression might be even larger if more listeners would be included in the N_4 group. As mentioned earlier, for those listeners, the operational point is shifted toward higher SNRs relative to the N_2 and N_3 groups, and hence the *acoustical* differences between the strategies are greater. It is even possible that at such high SNRs, the differences in perception are driven mostly by the changes in the acoustics of speech, that is, the high ECR of speech achieved by conventional fast-acting and SNR-aware processing compared with slow-acting compression (see May et al., 2018; Figure 3), and not by the interaction of speech and noise. If this was the case, the speech-intelligibility benefit of fast- over slow-acting compression would increase with increasing SRT. However, a regression analysis did not indicate any significant correlation of the two outcomes. Moreover, this

prediction is based on an assumption that applying fast-acting compression to the target is always desirable. It is possible that, due to greater suprathreshold auditory processing deficits, more severely impaired listeners rely more strongly on the temporal-envelope cues present in the speech signal itself—a notion supported by the studies of Souza et al. (2005), Souza et al. (2012b), and Souza et al. (2015b). In that case, *any* form of fast-acting compression could be detrimental to speech recognition by those listeners, negatively affecting their perception despite seemingly positive acoustical effects. The regression analysis revealed that neither the pure-tone average nor age could predict the differences in performance between processing types. Some form of a psychoacoustic metric of sensitivity to temporal-envelope distortion could potentially identify the listeners who are likely to be negatively affected by fast-acting processing. However, to date no such test exists, especially taking practical considerations in a clinical environment into account. Some evidence suggests that HI individuals with high working-memory capacity are better able to take advantage of fast-acting processing of the speech signal (see Souza et al. 2015a, for a review). It is possible that, in this study, such participants took greater advantage of the differential processing of the target and the background noise. A measure of working-memory capacity was not included in this study design. Nevertheless, considering this factor in future investigations could help to establish whether the cognitively high-performing listeners indeed benefit more from the SNR-aware compression strategies and hence allow for a more individualized fitting.

Limitations

The paired comparisons were conducted using noisy speech at a relatively low SNR, corresponding to the SRT. This allowed to measure both intelligibility and subjective preference in the same acoustic conditions. However, such conditions are not optimal for evaluating the overall sound quality, because listeners may not be able to focus on a broader range of attributes due to the low intelligibility. The listeners' preference might, in fact, be confounded solely by the differences in intelligibility between the processing types. A potential solution would be to adjust the SNR individually for each processing type, that is, to measure the SRT for all processing types instead of measuring it only for the fast-acting compression, reflecting an *iso-intelligibility* rather than an *iso-SNR* comparison. One could also conduct the paired comparisons at a higher SNR or even at a range of SNRs, revealing any potential effects of the SNR on the subjective preference. Moreover, apart from the overall preference, an explicit evaluation in terms of specific attributes such as subjective intelligibility, noisiness,

or clarity could be employed, as was done in the studies of Neuman et al. (1998) and Hansen (2002).

The frequency response of the headphone was equalized to have a flat response with reference to the ear-canal entrance, as described in the *Stimuli and procedure* subsection. As a consequence, the acoustic gain due to the pinna and the concha was not included in the presentation. This reduced the unaided response by 5 to 10 dB in the 2 to 4 kHz range, leading to a systematic mismatch between the aided response the NAL-NL2 target. Nevertheless, an exact match to the NAL-NL2 target for each listener was only possible at relatively low SPLs. This is because the level- and frequency-dependent gain values were based on the individual targets only for input SPLs up to 50 dB. At higher input SPLs, the gain was based on CRs that were fitted on a group level (N_2 , N_3 , and N_4) and not on the individual prescription. Moreover, this mismatch is mostly within the fit-to-target tolerances of ± 5 dB for frequencies up to 2 kHz and ± 8 dB above 2 kHz, as recommended by Gatehouse et al. (2001) and widely used in clinical settings during real-ear verification. While this effect might have affected audibility and spectral shaping (potentially relevant for sound quality), it has been present across all compression settings and was included in the SRT determination. Therefore, it is unlikely that it would have affected the study outcomes.

Finally, the results presented in this study evaluated the ideal SNR-aware compression strategy based on the *a priori* SNR. To apply this strategy in the context of hearing aids, the ideal speech detector needs to be replaced by an estimator that only has access to the noisy speech signal. The comparison in May et al. (2018) showed that a set of instrumental metrics was very similar for the SNR-aware system using either the estimated or the *a priori* SNR, indicating that a similar performance may be expected in the perceptual tasks. However, future work should evaluate the influence of SNR estimation errors on perception via behavioral listening tests.

Applicability to Real-World Scenarios

This study focused on the perceptual benefit of SNR-aware compression when processing noisy speech. This study did not take the effect of the overall SPL of the speech and noise components into account. The conditions were chosen to emphasize the influence of audibility on the outcome metrics; hence, a relatively low input noise SPL of 50 dB was selected. Hence, in many cases, the speech level was below normal conversational levels. It is possible that the balance between different cues provided by slow- and fast-acting compression would change at higher noise levels, which occur quite frequently in real-world scenarios (Smeds et al., 2015;

Weisser & Buchholz, 2019). It would therefore be interesting to investigate a condition with a notably higher background noise SPL (i.e., 65 or 70 dB).

Another factor that is present in many real-world acoustic scenarios, but not considered here, is reverberation. To take advantage of fast-acting compression of the speech signal in even more realistic scenarios where both room reverberation and interfering noise are present simultaneously, it is necessary to update the speech detection stage (e.g., with the power spectral density estimator proposed by Kuklasinski et al., 2016). When dealing with multiple competing sound sources that are spatially separated, the detection of speech-dominated T-F units could alternatively be accomplished by the analysis of spatial cues (May et al., 2011).

Conclusion

A perceptual evaluation of the SNR-aware compression strategy proposed by May et al. (2018) was conducted in controlled laboratory conditions in a group of HI listeners. The strategy was shown to provide a speech intelligibility benefit in noise compared with conventional slow-acting compression and achieved a higher subjective preference compared with both conventional fast- and slow-acting compression schemes. Future research will characterize those listeners that benefit the most from this new compression scheme and will determine the applicability to a broader range of acoustic conditions.



Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported by the Technical University of Denmark and Centre for Applied Hearing Research.

ORCID iDs

Borys Kowalewski  <https://orcid.org/0000-0001-6385-6154>
Tobias May  <https://orcid.org/0000-0002-5463-5509>

Notes

1. Two definitions of the ECR have commonly been used in previous work. One is based on the reduction of modulation depth as a function of the modulation rate and has been used by Barfod (1978), Braida et al. (1982), and Stone and Moore (1992), among others. Another definition is based on the reduction of the dynamic range derived from short-term level histograms before and after processing, as used by

- Souza et al. (2006), Henning and Bentler (2017), Alexander and Rallapalli (2017), and May et al. (2018). In this study, the histogram-based approach was chosen as it tends to decouple the two effects that compression has on the speech signal: (a) increased audibility of the low-level components and (b) distortion of the temporal envelope and the associated cues for intelligibility. The metric of the temporal-envelope distortion as a function of modulation rate is referred to as the MTF.
- The Brüel & Kjær Type 4153 ear simulator (designed according to IEC, 2009) loads the earphone with an acoustic impedance similar to that of the human ear. The response measured via the in-built microphone approximates the SPL as a function of frequency at the entrance of the ear canal (not at the eardrum, as would be the case, e.g., for the commonly used 711 coupler). Therefore, the individual ear-canal response was superimposed on the signal when the headphones were coupled to the listeners' ears. Ćirić and Hammershøi (2006) measured the response of HDA200 headphones (the same as used in this study) coupled to the Type 4153 simulator and to 34 human ears (measured at the entrance to the ear canal) and developed an average ear-to-coupler difference response. This response (compared with the averaged free-field-to-ear-canal-entrance response) de-emphasizes the 2 to 4 kHz range by 5 to 10 dB, as it does not include the acoustic gain from the pinna and concha. When signals are delivered through headphones equalized to a flat response in the 4153 simulator, the unaided response (and consequently the aided response) at the eardrum will lack this gain component and will be below the target prescribed by the NAL-NL2 rationale.
 - Trials were assumed to be consistent if the preferences were transitive, that is, if item X was preferred over item Y and Y over Z , then it was expected that X would also be preferred over Z . If this was not the case and Z was instead preferred over X , the preferences were arranged in a so-called *circular triad*. In such cases, the items could not be arranged unambiguously into a ranking, which was a requirement for the subsequent data analysis. Such inconsistent trials were assumed to result from listener errors, loss of focus, or misunderstanding of the task and were therefore rejected.
 - Stochastic transitivity describes overall consistency of a group of listeners based on the pooled response matrix. It refers to the relative frequencies of the preference of one item over another. For example, if the frequency of preference of item X over item Y exceeded 0.5, and the frequency of preference of item Y over item Z also exceeded 0.5, then the frequency of preference of Z over X should not exceed 0.5. Stochastic transitivity has to be fulfilled to fit the Bradley–Terry–Luce (BTL) model to the pooled response data.
- References**
- Alexander, J. M., & Rallapalli, V. (2017). Acoustic and perceptual effects of amplitude and frequency compression on high-frequency speech. *The Journal of the Acoustical Society of America*, *142*(2), 908–923. <https://doi.org/10.1121/1.4997938>
- Bacon, S. P., & Oxenham, A. J. (2004). Psychophysical manifestations of compression: Hearing-impaired listeners. In S. P. Bacon, R. R. Fay, & A. N. Popper (Eds.), *Compression: From cochlea to cochlear implants* (pp. 107–152) Springer. https://doi.org/10.1007/0-387-21530-1_4
- Barfod, J. (1978). Automatic regulation systems with relevance to hearing aids. *Scandinavian Audiology. Supplementum*, *6*, 355–378.
- Bisgaard, N., Vlaming, M. S., & Dahlquist, M. (2010). Standard audiograms for the IEC 60118-15 measurement procedure. *Trends in Amplification*, *14*(2), 113–120. <https://doi.org/10.1177/1084713810379609>
- Bradley, R. A., & Terry, M. E. (1952). Rank analysis of incomplete block designs: I. The method of paired comparisons. *Biometrika*, *39*(3/4), 324–345. <https://doi.org/10.2307/2334029>
- Braida, L. D., Durlach, N. I., De Gennaro, S. V., Peterson, P. M., & Bustamante, D. K. (1982). Review of recent research on multiband amplitude compression for the hearing impaired. In G. A. Studebaker & F. H. Bess (Eds.), *Monographs in contemporary audiology* (pp. 133–140). The Vanderbilt Hearing-Aid Report.
- Ćirić, D., & Hammershøi, D. (2006). Coupling of earphones to human ears and to standard coupler. *The Journal of the Acoustical Society of America*, *120*(4), 2096–2107. <https://doi.org/10.1121/1.2258929>
- Cohen, J. (1973). Eta-squared and partial eta-squared in fixed factor ANOVA designs. *Educational and Psychological Measurement*, *33*(1), 107–112. <https://doi.org/10.1177/001316447303300111>
- Davies-Venn, E., Souza, P., Brennan, M., & Stecker, G. C. (2009). Effects of audibility and multichannel wide dynamic range compression on consonant recognition for listeners with severe hearing loss. *Ear and Hearing*, *30*(5), 494–504. <https://doi.org/10.1097/AUD.0b013e3181aacc5bc>
- Desloge, J. G., Reed, C. M., Braida, L. D., Perez, Z. D., & D'Aquila, L. A. (2017). Masking release for hearing-impaired listeners: The effect of increased audibility through reduction of amplitude variability. *The Journal of the Acoustical Society of America*, *141*(6), 4452–4465. <https://doi.org/10.1121/1.4985186>
- Dreschler, W. A., Verschuure, H., Ludvigsen, C., & Westermann, S. (2001). ICRA noises: Artificial noise signals with speech-like spectral and temporal properties for hearing instrument assessment. *International Journal of Audiology*, *40*(3), 148–157. <https://doi.org/10.3109/00206090109073110>
- Edwards, B. (2004). Hearing aids and hearing impairment. In S. Greenberg, W. A. Ainsworth, & R. R. Fay (Eds.), *Speech processing in the auditory system* (Chap. 7, pp. 339–421). Springer. https://doi.org/10.1007/0-387-21575-1_7
- Elberling, C., Ludvigsen, C., & Lyregaard, P. (1989). Dantale: A new Danish speech material. *Scandinavian Audiology*, *18*(3), 169–175. <https://doi.org/10.3109/01050398909070742>
- Ellermeier, W., Mader, M., & Daniel, P. (2004). Scaling the unpleasantness of sounds according to the BTL model: Ratio-scale representation and psychoacoustical analysis. *Acta Acustica United with Acustica*, *90*(1), 101–107.
- Gatehouse, S., Naylor, G., & Elberling, C. (2006). Linear and nonlinear hearing aid fittings—2. Patterns of candidature.

- International Journal of Audiology*, 45(3), 153–171. <https://doi.org/10.1080/14992020500429484>
- Gatehouse, S., Stephens, S. D. G., Davis, A. C., & Bamford, J. M. (2001). Good practice guidance for adult hearing aid fittings and services. *British Association of Audiological Scientists Newsletter*, 36.
- Giannoulis, D., Massberg, M., & Reiss, J. D. (2012). Digital dynamic range compressor design—A tutorial and analysis. *Journal of the Audio Engineering Society*, 60(6), 399–408.
- Hagerman, B., & Olofsson, Å. (2004). A method to measure the effect of noise reduction algorithms using simultaneous speech and noise. *Acta Acustica United with Acustica*, 90(2), 356–361.
- Hansen, M. (2002). Effects of multi-channel compression time constants on subjectively perceived sound quality and speech intelligibility. *Ear and Hearing*, 23(4), 369–380.
- Hassager, H. G., May, T., Wiinberg, A., & Dau, T. (2017). Preserving spatial perception in rooms using direct-sound driven dynamic range compression. *The Journal of the Acoustical Society of America*, 141(6), 4556–4566. <https://doi.org/10.1121/1.4984040>
- Henning, R. L. W., & Bentler, R. A. (2017). The effects of hearing aid compression parameters on the short-term dynamic range of continuous speech. *Journal of Speech, Language, and Hearing Research*, 51(2), 471–484. [https://doi.org/10.1044/1092-4388\(2008/034\)](https://doi.org/10.1044/1092-4388(2008/034))
- International Electrotechnical Commission. (2009). *IEC 60318-1. Simulators of human head and ear. Part 1: Ear Simulator for the measurement of supra-aural and circum-aural earphones*.
- Jenstad, L. M., & Souza, P. E. (2005). Quantifying the effect of compression hearing aid release time on speech acoustics and intelligibility. *Journal of Speech, Language, and Hearing Research*, 48(3), 651–667. [https://doi.org/10.1044/1092-4388\(2005/045\)](https://doi.org/10.1044/1092-4388(2005/045))
- Jenstad, L. M., & Souza, P. E. (2007). Temporal envelope changes of compression and speech rate: Combined effects on recognition for older adults. *Journal of Speech, Language, and Hearing Research*, 50(5), 1123–1138. [https://doi.org/10.1044/1092-4388\(2007/078\)](https://doi.org/10.1044/1092-4388(2007/078))
- Kates, J. M. (1993). Optimal estimation of hearing-aid compression parameters. *The Journal of the Acoustical Society of America*, 94(1), 1–12. <https://doi.org/10.1121/1.407078>
- Keidser, G., Dillon, H., Flax, M., Ching, T., & Brewer, S. (2011). The NAL-NL2 prescription procedure. *Audiology Research*, 1(e24), 88–90. <https://doi.org/10.4081/audiores.2011.e24>
- Kendall, M. G. (1962). Paired comparisons. In *Rank correlation methods* (3rd ed., Chap. 11, pp. 144–154). Charles Griffin & Company Limited.
- Kendall, M. G., & Smith, B. B. (1940). On the method of paired comparisons. *Biometrika*, 31(3/4), 324–345. <https://doi.org/10.2307/2332613>
- Killion, M. C., Teder, H., Johnson, A. C., & Hanke, S. P. (1992). *Variable recovery time circuit for use with wide dynamic range automatic gain control for hearing aid* (U. S. Patent No. 5,144,675).
- Kowalewski, B., Zaar, J., Fereczkowski, M., MacDonald, E. N., Strelcyk, O., May, T., & Dau, T. (2018). Effects of slow- and fast-acting compression on hearing-impaired listeners' consonant-vowel identification in interrupted noise. *Trends in Hearing*, 22. <https://doi.org/10.1177/2331216518800870>
- Kuklasinski, A., Doclo, S., Jensen, S. H., & Jensen, J. (2016). Maximum likelihood PSD estimation for speech enhancement in reverberation and noise. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 24(9), 1599–1612. <https://doi.org/10.1109/TASLP.2016.2573591>
- Lai, Y. H., Li, P. C., Tsai, K. S., Chu, W. C., & Young, S. T. (2013). Measuring the long-term SNRs of static and adaptive compression amplification techniques for speech in noise. *Journal of the American Academy of Audiology*, 24(8), 671–683. <https://doi.org/10.3766/jaaa.24.8.4>
- Luce, R. D. (1959). On the possible psychophysical laws. *Psychological Review*, 66(2), 81–95. <https://doi.org/10.1037/h0043178>
- May, T., Kowalewski, B., & Dau, T. (2018). Signal-to-noise-ratio-aware dynamic range compression in hearing aids. *Trends in Hearing* 22, 1–12. <https://doi.org/10.1177/2331216518790903>
- May, T., van de Par, S., & Kohlrausch, A. (2011). A probabilistic model for robust localization based on a binaural auditory front-end. *IEEE Transactions on Audio, Speech, and Language Processing*, 19(1), 1–13. <https://doi.org/10.1109/TASL.2010.2042128>
- Moore, B. C. J. (2008). The choice of compression speed in hearing aids: Theoretical and practical considerations and the role of individual differences. *Trends in Amplification*, 12(2), 103–112. <https://doi.org/10.1177/1084713808317819>
- Moore, B. C. J., & Glasberg, B. R. (1988). A comparison of four methods of implementing automatic gain control (AGC) in hearing aids. *British Journal of Audiology*, 22(2), 93–104. <https://doi.org/10.3109/03005368809077803>
- Naylor, G. (2016). Theoretical issues of validity in the measurement of aided speech reception threshold in noise for comparing nonlinear hearing aid systems. *Journal of the American Academy of Audiology*, 27(7), 504–514. <https://doi.org/10.3766/jaaa.15093>
- Naylor, G., & Johannesson, R. B. (2009). Long-term signal-to-noise ratio at the input and output of amplitude-compression systems. *Journal of the American Academy of Audiology*, 20(3), 161–171. <https://doi.org/10.3766/jaaa.20.3.2>
- Neuman, A. C., Bakke, M. H., Mackersie, C., Hellman, S., & Levitt, H. (1995). Effect of release time in compression hearing aids: Paired-comparison judgments of quality. *The Journal of the Acoustical Society of America*, 98(6), 3182–3187. <https://doi.org/10.1121/1.413807>
- Neuman, A. C., Bakke, M. H., Mackersie, C., Hellman, S., & Levitt, H. (1998). The effect of compression ratio and release time on the categorical rating of sound quality. *The Journal of the Acoustical Society of America*, 103(5), 2273–2281. <https://doi.org/10.1121/1.422745>
- Neumann, J. (2008). *Method for dynamic determination of time constants, method for level detection, method for compressing an electric audio signal and hearing aid, wherein the method for compression is used* (U. S. Patent No. 7,333,623).

- Nielsen, J. B., & Dau, T. (2011). The Danish hearing in noise test. *International Journal of Audiology*, *50*(3), 202–208. <https://doi.org/10.3109/14992027.2010.524254>
- Pavlovic, C. V., & Studebaker, G. A. (1984). An evaluation of some assumptions underlying the articulation index. *The Journal of the Acoustical Society of America*, *75*(5), 1606–1612. <https://doi.org/10.1121/1.390870>
- Plomp, R. (1988). The negative effect of amplitude compression in multichannel hearing aids in the light of the modulation-transfer function. *The Journal of the Acoustical Society of America*, *83*(6), 2322–2327. <https://doi.org/10.1121/1.396363>
- Preminger, J. E., & Van Tasell, D. J. (1995). Quantifying the relation between speech quality and speech intelligibility. *Journal of Speech, Language, and Hearing Research*, *38*(3), 714–725. <https://doi.org/10.1044/jshr.3803.714>
- Rhebergen, K. S., Maalderink, T. H., & Dreschler, W. A. (2017). Characterizing speech intelligibility in noise after wide dynamic range compression. *Ear and Hearing*, *38*(2), 194–204. <https://doi.org/10.1097/AUD.0000000000000369>
- Rhebergen, K. S., Versfeld, N. J., & Dreschler, W. A. (2009). The dynamic range of speech, compression, and its effect on the speech reception threshold in stationary and interrupted noise. *The Journal of the Acoustical Society of America*, *126*(6), 3236–3245. <https://doi.org/10.1121/1.3257225>
- Smeds, K., & Leijon, A. (2011). Loudness and hearing loss. In M. Florentine (Ed.), *Loudness* (pp. 223–259). Springer. https://doi.org/10.1007/978-1-4419-6712-1_9
- Smeds, K., Wolters, F., & Rung, M. (2015). Estimation of signal-to-noise ratios in realistic sound scenarios. *Journal of the American Academy of Audiology*, *26*(2), 183–196. <https://doi.org/10.3766/jaaa.26.2.7>
- Souza, P. E., Arehart, K., & Neher, T. (2015a). Working memory and hearing aid processing: Literature findings, future directions, and clinical applications. *Frontiers in Psychology*, *6*, 1894. <https://doi.org/10.3389/fpsyg.2015.01894>
- Souza, P. E., Hoover, E., & Gallun, F. (2012a). Application of the envelope difference index to spectrally sparse speech. *Journal of Speech, Language, and Hearing Research*, *55*(3), 824–837. [https://doi.org/10.1044/1092-4388\(2011\)10-0301](https://doi.org/10.1044/1092-4388(2011)10-0301)
- Souza, P. E., Wright, R., & Bor, S. (2012b). Consequences of broad auditory filters for identification of multichannel-compressed vowels. *Journal of Speech, Language, and Hearing Research*, *55*(2), 474–486. [https://doi.org/10.1044/1092-4388\(2011\)10-0238](https://doi.org/10.1044/1092-4388(2011)10-0238)
- Souza, P. E. (2002). Effects of compression on speech acoustics, intelligibility, and sound quality. *Trends in Amplification*, *6*(4), 131–165. <https://doi.org/10.1177/108471380200600402>
- Souza, P. E., & Bishop, R. D. (1999). Improving speech audibility with wide dynamic range compression in listeners with severe sensorineural loss. *Ear and Hearing*, *20*(6), 461–470.
- Souza, P. E., Jenstad, L. M., & Boike, K. T. (2006). Measuring the acoustic effects of compression amplification on speech in noise. *The Journal of the Acoustical Society of America*, *119*(1), 41–44. <https://doi.org/10.1121/1.2108861>
- Souza, P. E., Jenstad, L. M., & Folino, R. (2005). Using multichannel wide-dynamic range compression in severely hearing-impaired listeners: Effects on speech recognition and quality. *Ear and Hearing*, *26*(2), 120–131.
- Souza, P. E., & Turner, C. W. (1998). Multichannel compression, temporal cues, and audibility. *Journal of Speech, Language, and Hearing Research*, *41*(2), 315–326. <https://doi.org/10.1044/jslhr.4102.315>
- Souza, P. E., & Turner, C. W. (1999). Quantifying the contribution of audibility to recognition of compression-amplified speech. *Ear and Hearing*, *20*(1), 12–20.
- Souza, P. E., Wright, R. A., Blackburn, M. C., Tatman, R., & Gallun, F. J. (2015b). Individual sensitivity to spectral and temporal cues in listeners with hearing impairment. *Journal of Speech, Language, and Hearing Research*, *58*(2), 520–534. https://doi.org/10.1044/2015_JSLHR-H-14-0138
- Stone, M. A., & Moore, B. C. (1992). Syllabic compression: Effective compression ratios for signals modulated at different rates. *British Journal of Audiology*, *26*(6), 351–361. <https://doi.org/10.3109/03005369209076659>
- Stone, M. A., & Moore, B. C. (2003). Effect of the speed of a single-channel dynamic range compressor on intelligibility in a competing speech task. *The Journal of the Acoustical Society of America*, *114*(2), 1023–1034. <https://doi.org/10.1121/1.1592160>
- Stone, M. A., & Moore, B. C. (2004). Side effects of fast-acting dynamic range compression that affect intelligibility in a competing speech task. *The Journal of the Acoustical Society of America*, *116*(4), 2311–2323. <https://doi.org/10.1121/1.1784447>
- Stone, M. A., & Moore, B. C. (2007). Quantifying the effects of fast-acting compression on the envelope of speech. *The Journal of the Acoustical Society of America*, *121*(3), 1654–1664. <https://doi.org/10.1121/1.2434754>
- Stone, M. A., & Moore, B. C. (2008). Effects of spectrotemporal modulation changes produced by multi-channel compression on intelligibility in a competing-speech task. *The Journal of the Acoustical Society of America*, *123*(2), 1063–1076. <https://doi.org/10.1121/1.2821969>
- Studebaker, G. A. (1985). A “rationalized” arcsine transform. *Journal of Speech, Language, and Hearing Research*, *28*(3), 455–462. <https://doi.org/10.1044/jshr.2803.455>
- Varga, A. P., & Steeneken, H. J. M. (1993). Assessment for automatic speech recognition: II. NOISEX-92: A database and an experiment to study the effect of additive noise on speech recognition systems. *Speech Communication*, *12*(3), 247–251. [https://doi.org/10.1016/0167-6393\(93\)90095-3](https://doi.org/10.1016/0167-6393(93)90095-3)
- Villchur, E. (1973). Signal processing to improve speech intelligibility in perceptive deafness. *The Journal of the Acoustical Society of America*, *53*(6), 1646–1657. <https://doi.org/10.1121/1.1913514>
- Walaszek, J. (2008). *Effects of compression in hearing aids on the envelope of the speech signal, Signal based measures of the side-effects of the compression and their relation to speech intelligibility* [Master’s Thesis]. Technical University of Denmark, DTU, DK-2800 Kgs. Lyngby.
- Weisser, A., & Buchholz, J. M. (2019). Conversational speech levels and signal-to-noise ratios in realistic acoustic conditions. *The Journal of the Acoustical Society of America*, *145*(1), 349–360. <https://doi.org/10.1121/1.5087567>

- Wickelmaier, F., & Schmid, C. (2004). A Matlab function to estimate choice model parameters from paired-comparison data. *Behavior Research Methods, Instruments, & Computers*, 36(1), 29–40. <https://doi.org/10.3758/BF03195547>
- Woods, W. S., Kalluri, S., Pentony, S., & Nooraei, N. (2013). Predicting the effect of hearing loss and audibility on amplified speech reception in a multi-talker listening scenario. *The Journal of the Acoustical Society of America*, 133(6), 4268–4278. <https://doi.org/10.1121/1.4803859>
- Yund, E. W., & Buckles, K. M. (1995). Enhanced speech perception at low signal-to-noise ratios with multichannel compression hearing aids. *The Journal of the Acoustical Society of America*, 97(2), 1224–1240. <https://doi.org/10.1121/1.412232>